地震前后基准站 GPS 数据异常提取算法 *

李 南[1]，孔祥增[2]，林 岭[2]

(1. 福建农林大学计算机与信息学院，福州 350000; 2. 福建师范大学 数学与信息学院，福州 350000)

摘 要：目前对地震前兆异常的研究主要集中在“热”和“电”等方面，很少涉及基准站的 GPS 数据。然而，已经有学者证明震中附近基准站的 GPS 时间序列坐标数据中也蕴含着大地震的前兆信息。针对 2001~2010 年间美国本土发生的具有代表性的三个地震进行了研究，首次将 Martingale 理论运用于 GPS 数据处理，提出一种异常提取算法，进而对地震前后，震中附近多个基准站的 GPS 数据进行分析。实验结果表明算法能够有效的反映大地震前后 GPS 数据中异常的变化趋势，为使用 GPS 数据对大地震进行预报提供了更多可能。

关键词：地震；Martingale 理论；异常提取

中图分类号：TP391 doi: 10.3969/j.issn.1001-3695.2017.09.0931

Detection of anomalies in GPS data before and after earthquakes

Li Nan[1], Kong Xiangzeng [2], Lin Ling [2]

(1. *College of Computer & Information Science Fujian Agriculture & Forestry University*, *Fuzhou* 350000, *China*; 2. *College of Mathematics & Information Fujian Normal University*, *Fuzhou* 35000, *China*)

Abstract: Most research studies on forecasting of earthquakes focus on thermal or electrical aspects. Scholars have proved GPS time series also contains precursory information for large earthquakes. For the problems mentioned before, three typical earthquakes hit America from 2001 to 2010 are investigated by using GPS data of several GPS reference stations near the epicenters. An algorithm based on Martingale theory is proposed to detect the anomalies in GPS data. The experimental results show that the proposed algorithm can effectively reflect the change process in GPS data before and after large earthquakes, which offers more possibilities for earthquake forecast from GPS data.

Key Words: earthquake; Martingale theory; anomaly detection

0 引言

地震是一种常见的自然灾害,它会对震区人民群众的生命和财产造成巨大的损害。在过去的数十年中，不少学者致力于研究地震前兆异常，以对地震未来的发展趋势进行预测，涉及的指标主要包括“热”[1]和“电”[2]等方面。例如，Kong 等人[3]使用 NOAA 卫星遥感射出长波辐射（outgoing long-wave radiation, OLR），探索汶川地震和庐山地震前 OLR 数据的异常。Han 等人[4]统计了 2001~2010 年日本 Kakioka 站台的超低频段电磁发射数据，试图找出数据中异常和地震之间的关联。Nemec 等人[5]使用统计方法研究地震前低频电磁波数据，发现在地震前短时间内固定频率电磁波会出现异常。Liu 等人[6]研究了台湾地区 1993~2004 年电离层总电子含量异常和地震的相关性。Kuo 等人[7]对日本 2011 Tohoku-Oki Mw 9.0 级地震前，电离层总电子含量的预测值和观察值进行了对比，以研究地震前电离层 TEC 异常。

研究表明，基准站的 GPS 时间序列坐标数据经过前期处理，可以求出每日的坐标位移解，能反映中、低频地形形变信息，并且具有比较好的稳定性[8]。Wang 等人[9]使用 Molchan 图表法，已经证明 GPS 形变中蕴含着地震的前兆信息。Sagiya 等人[10]以多个 GPS 基准站时序坐标数据为对象，来研究地震等其他地壳变化。Yeha 等[11]基于多个 GPS 基准站建立一个 GPS 网络，发现 GPS 时序坐标数据在 2013 年台湾南投和花莲地震前出现了一定程度的扰动。文献[12]中尝试使用一种非线性过滤器来提取 GPS 数据中可能与地震有关的异常信号。赵等[13]利用中国大陆几个邻近 GPS 基准站数据，研究了日本 2011 年发生的 9.0 级地震的震前变化，认为震前运动速度变慢的趋势可能是地震的前兆。因此，尽管将 GPS 数据作为地震前兆异常中指标的研究还是相对较少，对其进行分析，寻找其中蕴含的地震前兆异常信息具有可行性。然而，现有成果主要局限于地理学科，结果分析主要是基于人工观察，直接受到领域专家知识和经验的影响，对于震前相关异常的判定具有很大的不确定性。

基金项目：国家自然科学基金青年项目（41601477）；福建省引导性项目（2015Y0054）；福建省自然科学基金资助项目（2016J01280）

作者简介：李南（1987-），男，福建福州人，讲师，硕士，主要研究方向为数据挖掘、地震异常检测（binbanbinban@163.com）；孔祥增（1981-），男，实验师，博士，主要研究方向为数据挖掘、人工智能；林岭（1973-），女，讲师，硕士，主要研究方向为数据挖掘、软件工程.

数据挖掘的目的就是从海量的、不完全的、有噪声的数据中提取出隐含的、未知的但是又潜在有用的知识和信息的过程。当前，已有部分学者将数据挖掘中异常检测的方法运用于分析地震前不同指标的变化。例如，Marzocchi 等人[14]使用贝叶斯估值方法运用于分析地震数据，文献[15]中使用小波变换用于寻找地震前 OLR 数据的异常，文献[16]中使用模糊神经算法用于检测震前电磁异常，文献[17]中提出一种基于误差和关键点的地震前兆观测数据异常挖掘算法。

本文的贡献主要在于：目前涉及使用基准站的 GPS 数据来分析震前异常的研究还比较少，方法也主要是基于人工观察，对于异常的判定不确定性较大。因此，本文使用 Nevada Geodetic Laboratory 提供的 GPS 数据，基于 Martingale 理论[18]与 GPS 数据特点，提出了一种异常提取算法 ADM（anomaly detection based on martingale theory for GPS data）。针对发生在美国比较具有代表性的三个地震，运用该算法，对地震前后震中附近基准站的 GPS 时间序列数据进行了分析研究。结果不仅进一步证实了 GPS 数据中蕴藏着地震的前兆信息，而且为将数据挖掘算法运用于地理学科提供了借鉴。

1　相关数据

本文研究的对象是从 2001 年以来，发生在美国陆地上深度小于 60km 的地震中，震级最大的 Ms7.2 级地震，以及两个震中距离相对较近的地震。地震的相关信息来自美国地质调查局（USGS，https://earthquake.usgs.gov/earthquakes/），具体如表 1 所示。

表 1　地震相关信息

Date (y-m-d)	Latitude (°N)	Longitude (°E)	Depth (km)	Magnitude (JMA)
01-02-28	47.149	-122.727	51.80	6.8
08-02-21	41.144	-114.872	7.90	6.2
10-04-04	32.286	-115.295	9.99	7.2

使用的 GPS 数据是 GPS 基准站站点坐标，分为东西向、北南向和垂直向三个方向，来源于 Nevada Geodetic Laboratory 的数据共享服务网站（http://geodesy.unr.edu/），数据使用 IGS08 框架[19]。综合考虑站点位置以及 GPS 数据的完整性，实验中使用的 GPS 基准站的信息如表 2 所示。SEAT 站和 CLOV 站分别对应 2001-02-28 与 2008-02-21 这两个发生在不同时间，震中位置相对较近的地震，以分析不同时期，同一地理位置附近，地震前后基准站的 GPS 数据。MEXI、IID2 与 P500 这三个站点对应的是发生在 2010-04-04 的 M7.2 级地震，以研究对于同一地震，震中附近不同基准站的 GPS 数据是否出现相似的异常。

2　研究方法

以 $Data=\{d_1,d_2,\ldots,d_{t-1}\}$ 表示某基准站指定方向上的 GPS 历史数据，d_t 表示当前数据。显然，当地质板块的活动较为平稳时，$Data$ 中的样本分布应该维持在一个相对稳定的水平，样本之间具有某种相似的特征。因此，结合文献[3]，可以将 $Data\cup d_t$ 看成按照时间顺序，以数据流形式到来的，来自同一个聚类中的样本进行分析。

表 2　GPS 站点相关信息

Name	Latitude (°N)	Longitude (°E)	Height (m)
SEAT	47.654	-122.309	44.211
CLOV	40.558	-114.874	2084.911
MEXI	32.633	-115.476	-22.339
IID2	32.706	-115.032	9.792
P500	32.690	-115.300	-20.595

承上所述，基于最小化同一聚类中各样本差异原则[20]，设 C 为 $Data$ 的聚类中心，以式（1）计算：

$$C=\frac{1}{t-1}\sum_{i=1}^{t-1}d_i \tag{1}$$

那么 d_t 对于 C 的偏移程度可以用式（2）衡量：

$$s_t=s(Data,d_t)=\|d_t-C\| \tag{2}$$

其中 $\|\bullet\|$ 可以是任意的相异性度量函数，实验中使用欧式距离。

考虑 $d_1,d_2,\ldots,d_{t-1}$ 的分布情况，可以依据 $Data$ 中 d_t 近邻对象的个数来衡量其异常程度。因此，ADM 算法中用式（3）来衡量 d_t 所对应的当前指定方向上 GPS 数据的平稳程度：

$$p_t=\frac{\#\{j\mid s_j>s_t\}+\theta_t\#\{j\mid s_j=s_t\}}{t} \tag{3}$$

在式（3）中，θ_t 是 $(0,1]$ 内的随机数，$\#\{\}$ 是一个函数，返回满足给定条件的样本的数量，$j=1,2,\ldots,t$，s_t 采用式（2）计算。

从式（3）可以看出，$p_t\in(0,1]$ 且 p_t 越大，表明 d_t 越符合之前样本的分布情况，相应的当天数据出现异常的可能性越小。然而，GPS 数据中不可避免地会存在噪声（闪烁噪声以及随机漫步噪声等），某天相对较小的 p_t 值并不足以说明当前 GPS 数据整体出现了异常，需要通过使用 Martingale 理论[18]，综合考虑整个时间序列数据的变化趋势。

Martingale 理论[18]起源于赌博业和概率论，是最早的金融资产定价模型之一，现在已经被广泛运用于决策优化，投资分析等领域。为了真实反映 GPS 数据持续变化的情况，结合 Martingale 理论与 GPS 数据特点，所需要分析的每个数据 d_t 所对应的随机化的 Martingale 值（以下简称 M 值）以式（4）计算：

$$M_t=\prod_{i=1}^{t}(\varepsilon p_i^{\varepsilon-1})=M_{t-1}\cdot(\varepsilon p_t^{\varepsilon-1}) \tag{4}$$

根据文献[21]，取 $\varepsilon=0.82$ 并初始化 $M_0=1$。从式（4）中 M 值的计算过程可以看出，越大的 M 值表示之前一定时间内 GPS

数据的异常越明显。

聚类算法中，初始聚类中心的选择会对聚类结果造成较大的影响。类似的，$Data$ 中首个样本 d_1 位置的选取也会对异常的提取造成较大的影响。因此，算法中设置了一个初始聚类样本个数参数 num。在初始化 $C=\frac{1}{num}\sum_{k=1}^{num} d_{num}$ 以及 $s_i=\|d_i-C\|, i=1,2,\ldots,num$ 后，从 d_{num+1} 开始计算 M 值。直观地说，就是将所分析数据前 num 天的平均数据作为初始中心，并从 num 天后开始分析以减少结果的不确定性。

此外，由于大地震发生前后地壳运动相对剧烈，表现为 GPS 数据有可能在短时间内出现较大波动。结合式（4），此时 M 值会迅速不断增大到不可控的程度。为了避免这种情况的发生，算法中设置了另一个停止参数 h。也就是说，如果 $M_i > h$，那么从 d_{i+1} 开始重新开始算法。参数 num 与 h 对算法效果的影响将在实验部分进行进一步分析与验证。

综上所述，ADM 算法的具体步骤如下：

算法名称：ADM

输　入：某基准站指定方向上的 GPS 数据 $Data$，

停止参数 h，

初始聚类样本数 num

输　出：$Data$ 中每个样本对应的 M 值

Step.1 初始化样本中心：$C_{num}=\frac{1}{num}\sum_{i=1}^{num} d_i$

偏移程度：$s_i=\|d_i-C\|, i=1,2,\ldots,num$

M 值：$M_i=1, i=1,2,\ldots,num$

当前处理样本序号：$k=num+1$

循环停止条件：$isStop=false$

当前批次样本开始序号：$front=1$

// 如果检测到 $M_k > h$，那么从 d_{k+1} 开始

// 一个新的批次（即算法重新启动）

Step.2　while　$k \le |Data|$　&&　！$isStop$

2.1 使用 d_k、$C_{k\text{-}1}$ 与式（2），计算 s_k

2.2 使用 $s_{front}, s_{front+1},\ldots,s_k$ 与式（3），计算 p_k

2.3 使用 M_{k-1} 与式（4），计算 M_k

2.4 计算 $d_{front}, d_{front+1},\ldots,d_k$ 的聚类中心

$$C_k=\frac{1}{k-front}\sum_{i=front}^{k} d_k$$

2.5 使用 C_k、$d_{front}, d_{front+1},\ldots,d_k$ 与式（2），更新 $s_{front}, s_{front+1},\ldots,s_k$

2.6　$k++$

2.7　if　$M_{k\text{-}1} > h$

$isStop=true$

end if

end while

Step. 3　if　$isStop=true$　&&　$k+num+1 \le |Data|$

3.1　$isStop=false$

3.2　$C_{k+num}=\frac{1}{num}\sum_{i=k}^{k+num} d_i$

3.3　$M_i=1, i=k,k+1,\ldots,k+num$

3.4　$front=k$

3.5　$k=k+num+1$

3.6　转到 Step. 2　//开始一个新的批次

end if

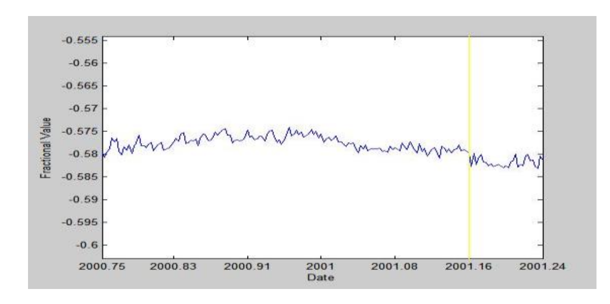
Step 4.　输出 $M_i, i=1,2,\ldots,|Data|$

3　结果与分析

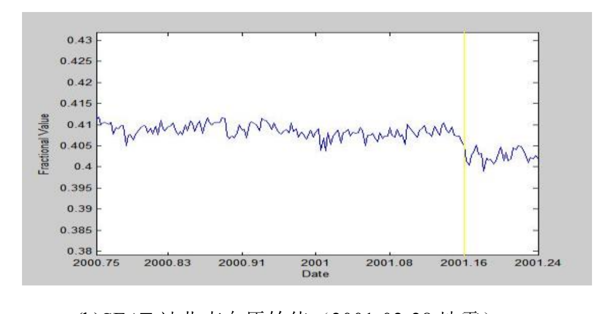
a）SEAT 站和 CLOV 站的原始 GPS 时间序列坐标数据以及对应的 M 值如图 1 和 2 所示。两个基准站分别对应两个发生在不同时间、震中位置相对较近的地震。ADM 算法中，参数 num 设置为 45，h 设置为 1000。研究的数据从发生地震上年的 10 月开始，一直延续到地震后的一个月。为了降低 θ 的随机取值对结果造成的影响，图中使用 10 次实验的平均值。出于更好的观察 GPS 数据变化趋势的目的，原始值中只使用相应 GPS 数值的小数部分而忽略相同的整数部分，黄色竖线表示地震发生时间。

从图 1（d）与图 2（d）可以看出，前期各个方向上的 M 值几乎没有任何变化，说明 GPS 数据比较稳定。临近地震前后，两个基准站东西向和北南向的 M 值均出现了显著变化，并且在地震后一周左右迅速达到峰值，这进一步证实了 GPS 数据中蕴含着和地震有关的信息。值得注意的是，M 值在地震前出现波动，在地震后才达到峰值的原因可能在于：由于地震的生成机制，GPS 数据的异常可能出现在地震前很短的一段时间内，而为了反映 GPS 数据的变化趋势以及减少噪声对结果造成的影响，算法需要在累积一定数量的异常样本之后，M 值才会相应出现显著的变化，因此地震前 GPS 数据发生的异常有可能相应的表现为地震后短时间内 M 值达到峰值。虽然可以通过将参数 h 设置为较小的数值以将 M 值峰值出现的时间提前，但是这有可能导致出现误报，本文将在实验 c）中验证参数 h 对算法效果的影响。

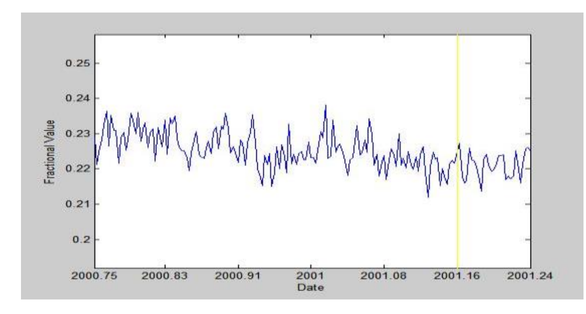
由于两个地震震中位置相对较近，成因相似，因此所提取的东西向和北南向的 M 值变化规律也比较类似，这从侧面上说明本文提出的 ADM 算法的可靠性。此外，这两个方向上的原始 GPS 数据图像并无出现明显的异常趋势，这证实了 ADM 算法提取 GPS 数据异常的有效性。因为垂直向 GPS 数值的测量误差要远远大于东西向和北南向[9]，所以对比东西向和北南向的 M 值，两个基准站垂直向 M 值波动都不明显，并没有呈现和地震明显相关的规律。



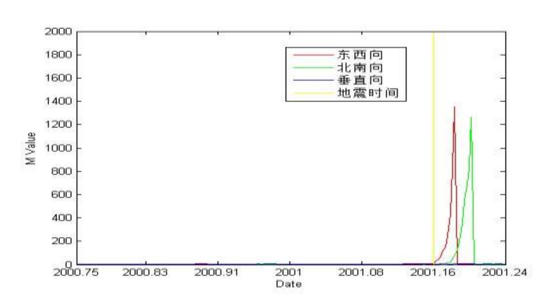
(a)SEAT 站东西向原始值（2001-02-28 地震）



(b)SEAT 站北南向原始值（2001-02-28 地震）



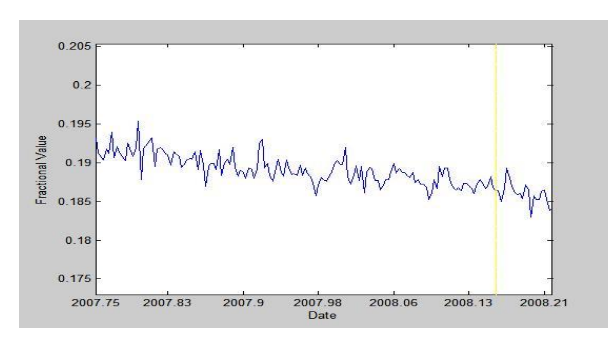
(c)SEAT 站垂直向原始值（2001-02-28 地震）



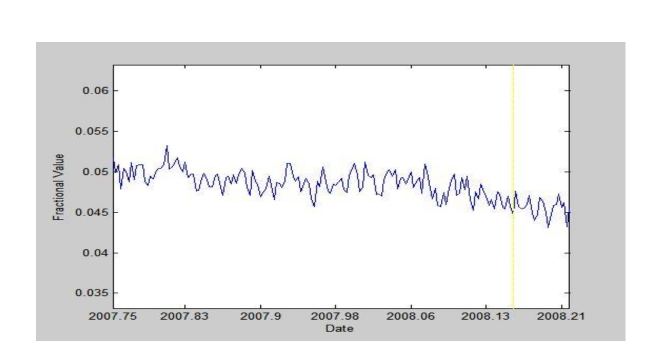
(d)SEAT 站三个方向对应 *M* 值（2001-02-28 地震）



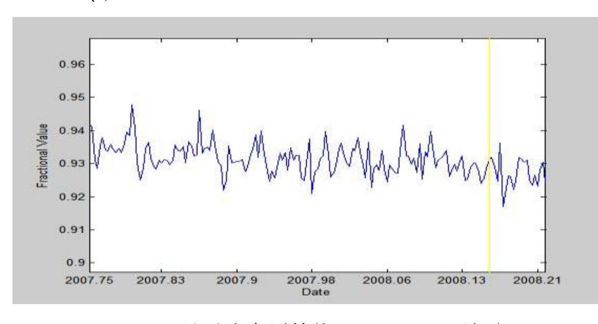
图 1



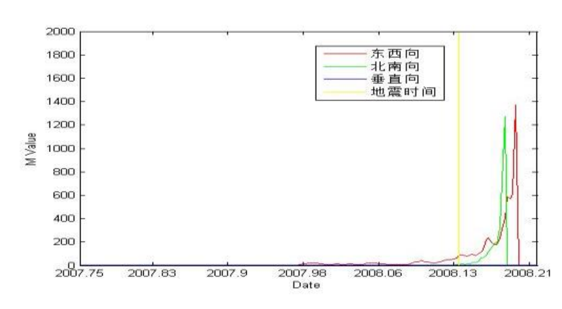
(a) CLOV 站东西向原始值（2008-02-21 地震）



(b)CLOV 站北南向原始值（2008-02-21 地震）



(c) CLOV 站垂直向原始值（2008-02-21 地震）



(d)CLOV 站三个方向 *M* 值（2008-02-21 地震）

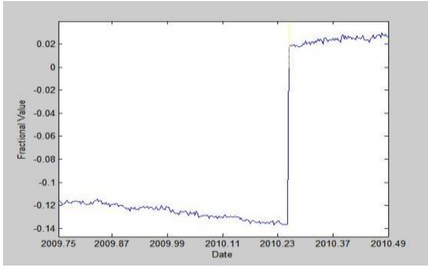
图 2

b）2010-04-04 地震震中邻近 MEXI 站、IID2 站和 P500 站的 GPS 数据原始值以及对应的 *M* 值分别见图 3、图 4 与图 5，原始值中同样只显示相应的小数部分，黄色竖线表示地震发生时间。实验的目的是研究对于同一地震，震中附近不同基准站 GPS 数据的异常。ADM 算法中，参数 *num* 设置为 45，*h* 设置为 1000。由于此次地震后短时间内，各个基准站特定方向上的 GPS 数据会突然出现“断崖”式的变化，因此研究的数据从发生地震上年的 10 月开始，一直延续到地震后的两个月。为了更清晰的展示地震前后，三个方向 GPS 数据的异常趋势，图 3~5 中分别单独列出对应三个方向上的 *M* 值。

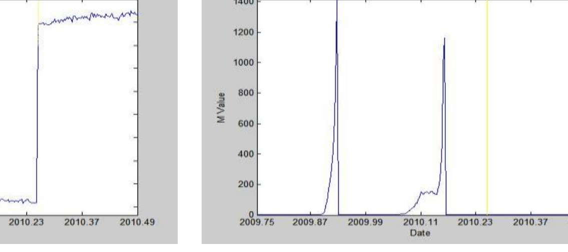
对比图 3(c) (d)可以看出，由于地震后基准站特定方向上的 GPS 数据出现的“断崖”式的变化，相应的 *M* 值也在之后对应的一段时间内达到峰值。类似的情况也出现在图 4(a)(b)，图 5(c)(d)等。波峰的延后出现，一方面说明本文提出的 ADM 算法能够真实反映 GPS 数据中的异常，另一方面也说明了实验 1 中导致 *M* 值出现峰值的 GPS 数据异常可能是在震前出现的。图 1(a)与图 1(b)中，MEXI 站东西向没有出现类似情况的原因在于：GPS 数据在地震前就出现了潜在的剧烈变化，导致 *M* 值超过了

预先设置的停止参数 $h=1000$，于是算法重新启动并拟合了这部分“断崖”数据。

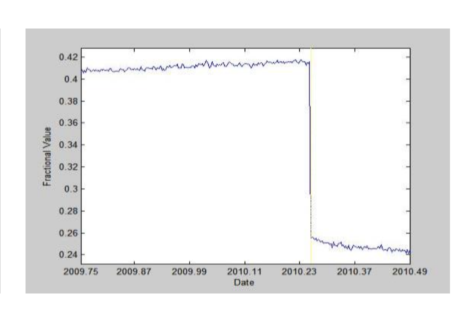
可能是由于地震成因不同并且 2010-04-04 地震的震级较大，因此三个基准站东西向的 M 值，MEXI 站与 IID2 站北南向的 M 值都在地震前一段时间就达到过峰值，说明在地震发生前 GPS 数据就有较大的异常发生，从侧面证明了 GPS 形变中蕴含着地震的前兆信息，尽管此时原始 GPS 数据图像并无明显的异常趋势。这为使用本文提出的 ADM 算法，利用 GPS 数据对大地震进行预报提供了可能。各个基准站第一次 M 值达到峰值后又出现平缓的原因在于：算法重新启动，需要拟合 num 天的数据后再判定异常，这段时间内 M 值始终为初始值 1。与 MEXI 站与 IID2 站类似，P500 站北南向的 M 值在地震前也出现了一定程度的变化，只不过变化程度要远低于上述两个基准站，这可能与基准站的具体位置有关。与实验 a）情况类似，三个基准站垂直向 M 值的变化并没有呈现一致的规律，只有 IID2 站垂直向的 M 值比较明显地在地震前后出现了波动，MEXI 站垂直向的波动可能来源于数据“断崖”式的变化，而 P500 站垂直向 M 值的峰值出现在地震后一段时间内。



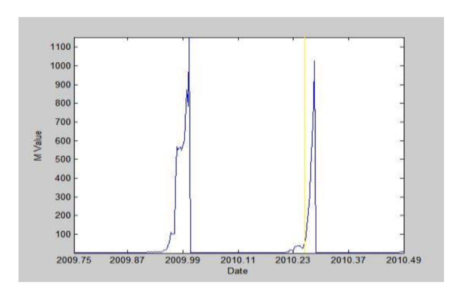
(a)MEXI 站东西向原始值(2010-04-04 地震)



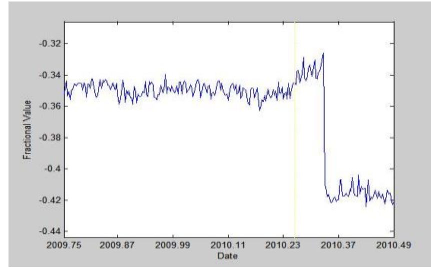
(b) MEXI 站东西向 M 值(2010-04-04 地震)



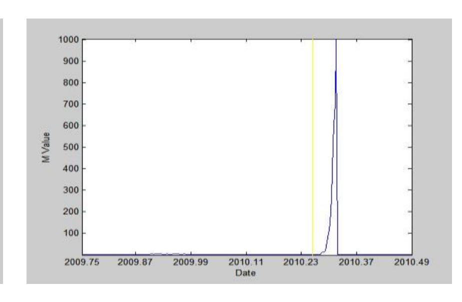
(c) MEXI 站北南向原始值(2010-04-04 地震)



(d) MEXI 站北南向 M 值(2010-04-04 地震)

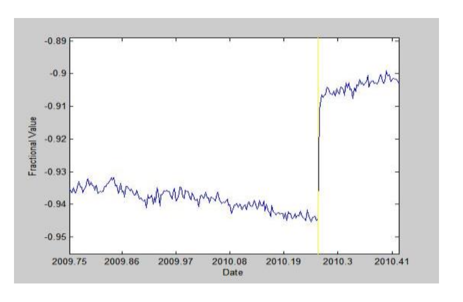


(e) MEXI 站垂直向原始值(2010-04-04 地震)

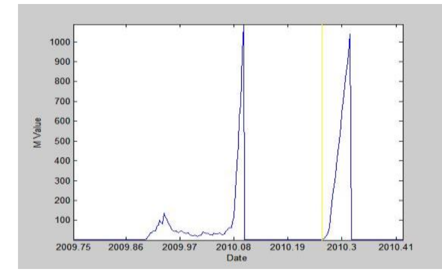


(f)MEXI 站垂直向 M 值(2010-04-04 地震)

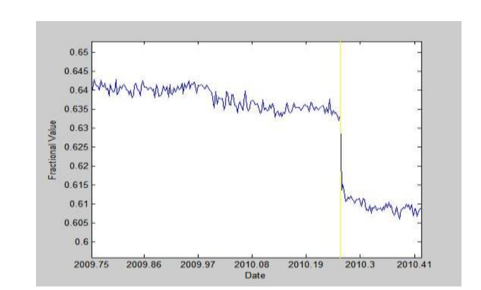
图 3



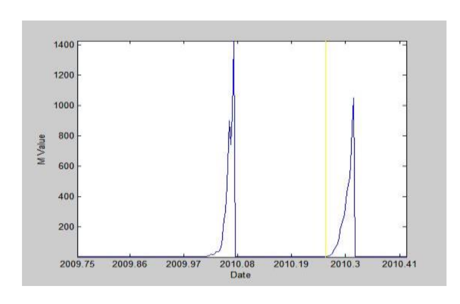
(a) IID2 站东西向原始值(2010-04-04 地震)



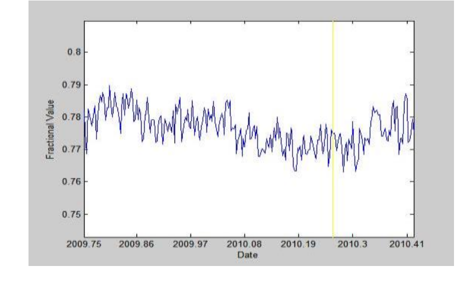
(b)IID2 站东西向 M 值(2010-04-04 地震



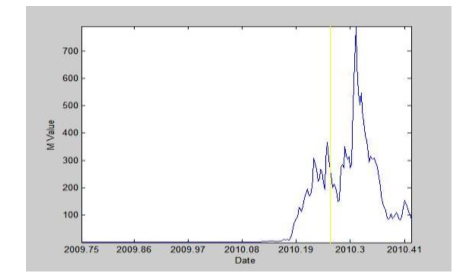
(c) IID2 站北南向原始值(2010-04-04 地震)



(d) IID2 站北南向 M 值(2010-04-04 地震)

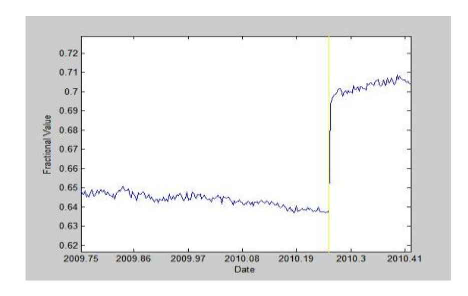


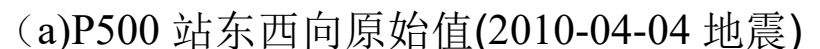
(e)IID2 站垂直向原始值(2010-04-04 地震)

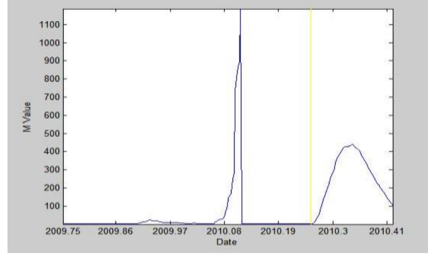


(f)IID2 站垂直向 M 值(2010-04-04 地震)

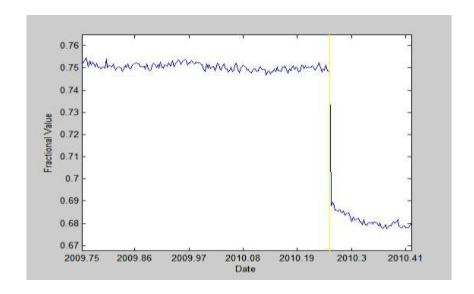
图 4



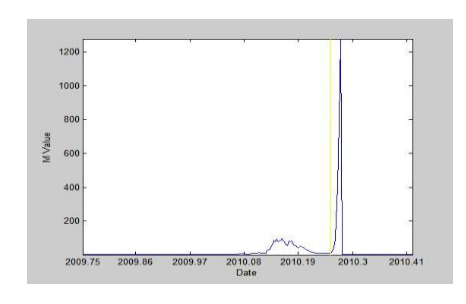

(a)P500 站东西向原始值(2010-04-04 地震)



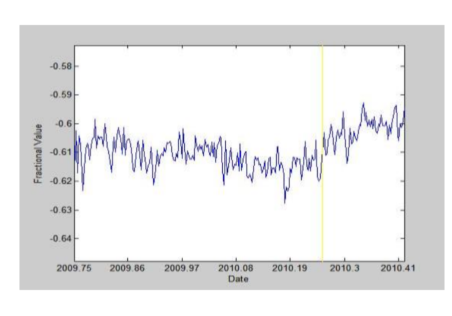
(b) P500 站东西向 M 值(2010-04-04 地震)



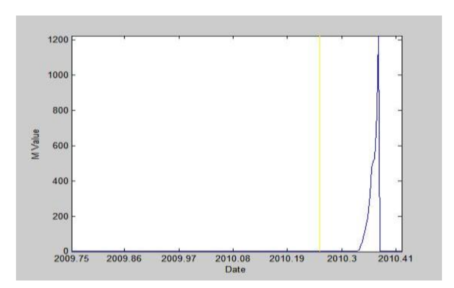
(c) P500 站北南向原始值(2010-04-04 地震)



(d) P500 站北南向 M 值(2010-04-04 地震)



(e) P500 站垂直向原始值(2010-04-04 地震)



(f) P500 站垂直向 M 值(2010-04-04 地震)

图 5

从基准站分布的位置上看，图 3~5 中可以看出，由于三个基准站都位于震中附近，因此在测量误差相对较小的东西向和北南向上，各个基准站 M 值第一次出现波峰的时间并不存在显著差异。但是，三个基准站中距离震中最近的 MEXI 站两个方向上波峰第一次出现的时间会早于 IID2 站与 P500 站。这说明距离震中越近的基准站，M 值可能越早出现波峰，即 GPS 数据潜在异常越早发生。

结合图 1~5，同样从测量误差相对较小的东西向和北南向上，进一步分析地震震级、发生时间与 M 值之间的关系。本文研究的三个地震中，震级最大的 2010-04-04 地震（Ms7.2）所对应的三个基准站的 M 值第一次波峰出现的时间均在地震发生前三个月左右，变化一直持续到地震发生时。而震级相对较小的 2001-02-28 地震（Ms6.8）与 2008-02-21 地震（Ms6.2）各自对应基准站的 M 值则最早在地震前一个月左右开始变化，并在地震后较短的一段时间内达到波峰。这说明地震震级越大，之前需要积攒的能量越多，GPS 数据潜在的异常可能更早发生，因此 M 值相应越早出现波动。此外，2001-02-28 地震（Ms6.8）M 值波动出现的时间要迟于 2008-02-21 地震（Ms6.2）的原因可能在于前者的震源深度（51.80km）要远大于后者（7.90km），因此地震发生前期 GPS 数据潜在变化相对不明显。

c）为了评价参数 h 对算法效果的影响，本文在 MEXI 站北南向上，分别将 h 设置为 250,500,1000,2000 进行了测试，参数 num 设置为 45。实验结果如图 6 所示。

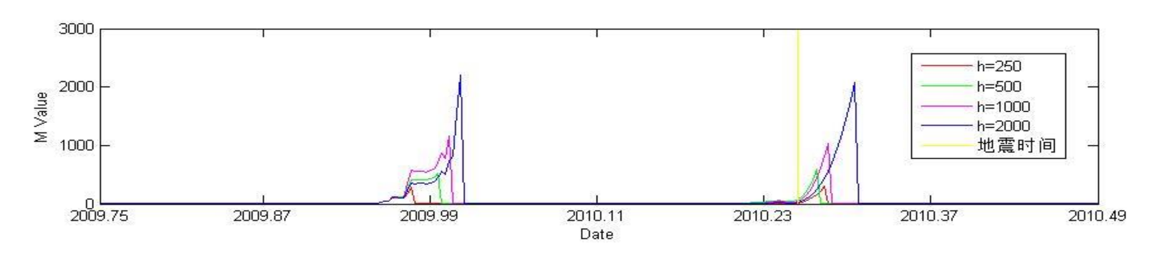


图 6　h 值对算法效果的影响(MEXI 站北南向)

从图 6 中可以看出，当 GPS 数据中出现异常时，M 值将以指数级别迅速增长。显然，当 num 相同时，通过将参数 h 设置为较小的数值，可以将首次 M 值峰值出现的时间提前。对比 500,1000,2000，将 h 设置为 250 时，第二次波峰出现的时间并没有提前的原因在于：当第一次检测到 $M \geq h$ 后，算法需要重新启动。此时，不同 h 值重新开始的位置不同，在 GPS 数据中已经出现异常的前提下，再次获得的新批次样本的初始聚类中心会有较大的差异，这有可能导致波峰出现滞后。由于过小的 h 值有可能会导致过多的误报，因此 $h=1000$ 是一个相对合理的取值。

d）为了评价参数 num 对算法效果的影响，本文同样在 MEXI 站北南向上，分别将 num 设置为 25,35,45,55 进行了测试，参数 h 设置为 1000。实验结果如图 7 所示。

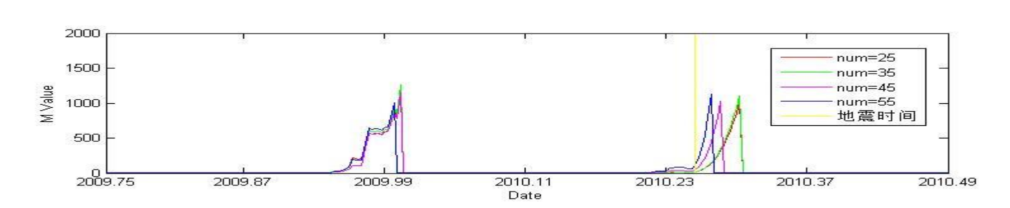


图 7　*num* 值对算法效果的影响(MEXI 站北南向)

显然，当 GPS 数据稳定时，只要 *num* 足够大，受其影响的，不同 *num* 值所对应的初始聚类中心 C 差别应该不大。因此，从图 7 中可以看出，*num* 值的不同对 M 值第一次波峰出现的时间几乎没有影响。然而，与实验（3）类似，当算法重新启动时，随着 GPS 数据中潜在异常的出现，不同 *num* 值所对应的初始聚类中心 C 有所差别，因此 M 值第二次出现波峰的时间存在一定的差异。虽然实验（4）中，将 *num* 设置为 55 时 M 值出现波峰的时间最早，但是考虑到将来实际应用中，过大的 *num* 值有可能使得 M 值两次波峰出现的间隔过长，因此本文中将 *num* 设置为 45。

e）为了比较本文提出的 ADM 算法与目前常用的利用领域专家知识和经验来判断地震前兆异常的方法[22]，本文在 SEAT 站北南向的数据上进行了实验，数据异常判定采用 $k\sigma$ 准则。在 $k\sigma$ 准则中，以 $\bar{x}$ 表示观测结果的平均值，σ 表示观测结果的标准差，如果数据超出了 $[\bar{x}-k\sigma,\bar{x}+k\sigma]$ 的范围，则认为有异常发生。实验中依据文献[22]将 k 值设定为 2，结果如图 8 所示。

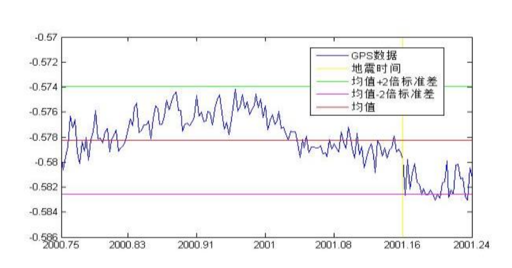


图 8　$k\sigma$ 准则实验结果(SEAT 站北南向)

从图 8 中可以看出，$k\sigma$ 准则不但没能及时发现震前异常，反而是地震发生数周后，数据多次超出了 $[\bar{x}-k\sigma,\bar{x}+k\sigma]$ 的范围，这表明 $k\sigma$ 准则并没有效果。对比图 1(d)中，地震前后短时间内，M 值不仅出现了波峰，同时数值能够反映大地震前后 GPS 数据中异常的变化趋势，显然本文提出的 ADM 算法更具有优势。

4　结束语

本文首次将 Martingale 理论运用于 GPS 数据，提出了一种针对 GPS 数据的异常提取算法 ADM，并用该算法分析地震前后，震中附近基准站的 GPS 时间序列坐标数据。从结果中可以看出，对比于局限于地理学科，基于领域专家知识和经验，直接观察数据的传统方法，M 值的变化能够真实反映 GPS 数据的潜在的变化趋势。即使原始 GPS 数值没有明显异常，在地震前后一段时间内其对应的 M 值也会出现较大的变化，而且有可能在地震前就达到峰值。这不仅进一步证实了 GPS 形变中蕴涵着地震的前兆信息，也为使用 GPS 数据对地震进行预报提供了可能。然而，基于目前的数据量和研究水平，探究使用基准站 GPS 数据对大地震进行预报将是一项长期工作。本文提出的 ADM 算法所提取的震中附近基准站 GPS 数据所对应的 M 值的异常已经揭示了一些初步规律，但还需要在更多的地震实例上进一步验证 M 值的变化趋势与地震各参数（震级、位置、发生时间等）之间的关系并对其进行统计可靠性分析，这也是本文下一步的研究方向。

参考文献：

[1] Saradjian M R, Alhoondzadeh M. Thermal anomalies detection before strong earthquakes (M>6, 0) using interquartile, wavelet and Kalman filter method [J]. Natural Hazards and Earth System Sciences, 2011, 11 (1): 1099-1108.

[2] Honkura Y, Oshiman N, Matsushima M, et al. Rapid changes in the electrical state of the 1999 izmit earthquake rupture zone [J]. Nature Communications, 2013, 4 (2116): 1-10.

[3] Kong X, Bi Y, Glass D H. Detecting seismic anomalies in outgoing long-wave radiation data [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8 (2): 649-660.

[4] Han P, Hattori K, Hirokawa M, et al. Statistical analysis of ULF seismomagnetic phenomena at Kakioka, Japan during 2001-2010 [J]. Journal of Geophysical Research: Space Physics, 2014, 119 (6): 4998-5011.

[5] Němec F, Santol K O, Parrot M. Decrease of intensity of ELF//VLF waves observed in the upper ionosphere close to earthquakes: A statistical study [J]. Journal of Geophysical Research Atmospheres, 2009, 114 (4): 1-10.

[6] Liu J Y, Chen Y I, Huang C H, et al. A statistical study of lightning activities and M ≥ 5. 0 earthquakes in Taiwan during 1993 – 2004 [J]. Surveys in Geophysics, 2015, 36 (6): 851-859.

[7] Kuo C L, Lee L C, Heki K. Preseismic TEC changes for Tohoku-Oki earthquake: Comparisons between simulations and observations. [J]. Terrestrial, Atmospheric & Oceanic Sciences, 2015, 26 (1): 63-72.

[8] 张燕，吴云，段维波，等. GPS 长趋势变化与大地震的关系 [J]. 武汉大学学报：信息科学版, 2012, 37 (6): 675-678.

[9] Wang T, Zhuang J, Kato T, et al. Assessing the potential improvement in short-term earthquaks forcasts from incorporation of GPS data [J]. Geophysical Research Letters, 2013, 40 (11): 1-5.

[10] Sagiya T, Miyazaki S, Tada T. Continuous GPS array and present-day crustal deformation of Japan [C]// Proc of the 1st Workshop of the APEC

Cooperation for Earthquake Simulation. [S. l.] : IEEE Press, 2000: 2303-2322.

[11] Yeha Y L, Chenga K C, Wanga W H, et al. Very short-term earthquake precursors from GPS signal interference based on the 2013 Nantou and Rueisuei earthquakes, Taiwan [J]. Journal of Asian Earth Sciences, 2015, 114 (2): 312-320.

[12] Wang T, Bebbington M. Identifying anomalous signals in GPS data using HMMs: An increased likelihood of earthquakes [J]. Comptutational Statistics and Data Analysis, 2013, 58 (1): 27-44.

[13] 赵国强，李鹏. 中国大陆 GPS 连续观测站测得的日本 9. 0 级地震同震位移与震前变化 [J]. 地震, 2012, 32 (2): 129-134.

[14] Marzocchi W, Zechar J, Jordan T. Bayesian forecast evalution and ensemble earthquake forcasting [J]. Bulletin of the Seismological Society of America, 2012, 102 (6): 2574-2584.

[15] Xiong P, Shen X, Bi Y, et al. Study of outgoing longwave radiation anomalies associated with Haiti earthquake [J]. Natural Hazards & Earth System Sciences, 2010, 10 (10): 2169-2178.

[16] Konstantaras A, Varley M, Vallianatos F, et al. Detection of weak seismo-electric signals upon the recordings of the electrotelluric field by means of neuro-fuzzy technology [J]. IEEE Geoscience & Remote Sensing Letters, 2007, 4 (1): 161-165.

[17] 李正媛，陈晶，王丽娜，等. 一种基于误差和关键点的地震前兆观测数据异常挖掘算法 [J]. 计算机应用研究, 2011, 28 (8): 2987-2901.

[18] Ho S S, Wechsler H. A Martingale framework for detecting changes in data streams by testing exchangeability [J]. IEEE Transactions on Software Engineering, 2010, 32 (12): 2113-2127.

[19] Rebischung P, Griffiths J, Ray J, et al. IGS08: the IGS realization of ITRF2008 [J]. GPS Solutions, 2012, 16 (4): 483-494.

[20] Kanungo T, Mount D M, S. Netanyahu N, et al. An efficient k-means clustering algorithm: Analysis and inplemention [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2002, 24 (7): 881-892.

[21] Vork V, Nouretdinov I, Gammerman A. Testing exchangeability on-line [C]// Proc of the 12th International Conference on Machine Learning, San Francisco: Morgan Kaufmann, 2003: 768-775.

[22] Wu L X, Qin K, Liu S J. GEOSS-Based thermal parameters analysis for earthquake anomaly recognition [J]. Proceedings of the IEEE, 2012, 100 (10): 2891-2907.